

Impact of Packetization and Functional Split on C-RAN Fronthaul Performance

Chia-Yu Chang, Ruggero Schiavi, Navid Nikaein, Thrasyvoulos Spyropoulos, Christian Bonnet
Mobile Communications Department
EURECOM, Biot, France 06410
Email: firstname.lastname@eurecom.fr

Abstract—Cloud-RAN (CRAN) is considered as one key enabler for beyond 4G networks, offering multiplexing gains, and advanced cooperation and coordinated signal processing. However, a key obstacle in the adoption of the CRAN architecture is that it requires very high capacity and low latency fronthaul (FH) links to carry raw I/Q samples between remote radio heads (RRH) and the baseband units (BBUs). These capacity requirements could be reduced by a more flexible split of baseband processing between BBUs and RRHs. Nevertheless, while moving some of the processing back into the RRH is expected to reduce FH rates, the amount of reduction mainly depends on the split, cell load, scenario and it might also introduce some delays. To this end, this paper studies the impact of different functional splits on the FH capacity for representative scenarios. Furthermore, we propose the use of a packet-based fronthaul network and study the joint impact of different packetization methods and RRH-BBU functional splits on the FH rate and latency. Based on this study, we provide some insights on the feasibility and optimality of different combinations, and the potential multiplexing benefits in terms of numbers of RRHs one could support over a single Ethernet-based FH network.

I. INTRODUCTION

The Cloud RAN or Centralized RAN (C-RAN) architecture is one of the most promising technologies that will impact future 5G architecture and possibly re-shape existing mobile network environments. Unlike traditional RAN, C-RAN detaches the Baseband units (BBU) from the edge radio equipments (the eNodeB). The baseband processing for many eNB, now called Remote Radio Heads (RRH), is centralized into a single pool of shared, and dynamically allocated BBUs, offering energy and multiplexing gains. These BBU functions could be implemented on commodity hardware and performed on virtual machines, further benefiting from softwarization and Network Function Virtualization (NFV). Finally, the centralization of BBU functions facilitates advanced coordinated multi-cell signal processing, which are often impractical in regular, distributed BS setups due to stringent synchronization constraints. An overview of Cloud RAN technology is provided in [1].

Despite its appeal, a key obstacle in the adoption of the C-RAN architecture is the excessive capacity and latency requirements on the Fronthaul (FH) link connecting an RRH with the BBU cloud. A simple example is depicted in Fig. 1): shifting all baseband processing away from the BS to a remote Cloud implies that to support a mere 75Mbps radio access rate, for a single category 5 single user, we need to transport 1Gbps of information on the FH link. If one further considers MIMO layers or carrier aggregation, these rates quickly

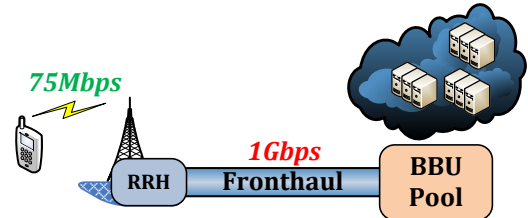


Fig. 1: FH data rate explosion

become prohibitive. Furthermore, the user expects to receive an ACK/NACK response within 4ms in maximum [2] after its transmission, imposing also a strong latency requirement on the FH link.

In order to relax these excessive FH bandwidth constraints, the operators and vendors are revisiting the concept of C-RAN, considering a more flexible distribution of baseband functionality between the RRH and the BBU pool [3]. Rather than offloading all the BBU processing on the cloud, dividing the Physical RX and TX chain in different blocks, it is possible to keep a subset of these blocks in the RRH. This concept is also known as *Flexible Centralization*. By gradually placing more and more BBU processing at the edge of the network, the FH capacity requirement becomes smaller (e.g., CRC and guard bands removed, Resource Elements that are idle are not sent, etc.). Nevertheless, flexible centralization has two main drawbacks, both relating to the initially envisioned benefits of C-RAN:

- (i) RRHs become more complex, and thus more expensive.
- (ii) De-centralizing the BBU processing reduces the opportunities for multiplexing gains, coordinated signal processing and advanced interference avoidance schemes.

Consequently, flexible or partial centralization is a trade-off between what is gained in terms FH requirements and what is lost in terms of C-RAN features.

Another key question is how the information between the RRH and the BBU is transported over the FH link. A number of FH transmission protocols are under investigation, such as the Common Public Radio Interface (CPRI) [4], CPRI 2 and OBSAI [5]. However, these have mainly been considered for carrying raw I/Q samples in a traditional C-RAN architecture. In light of the different possible functional splits, different types of information might need to be transported over the FH link. Given the extensive adoption of Ethernet in clouds, data centers, and the core network, Radio over Ethernet [6] could be a generic, cost-effective, off-the-shelf alternative for

FH transport. Furthermore, while a single FH link per RRH, all the way to the BBU pool, has usually been assumed, it is expected that the FH network will evolve to more complex multihop topologies, requiring switching and aggregation [7], such as the one depicted in Fig. 2. This is further facilitated by a standard Ethernet approach.

Nevertheless, packetization over the FH introduces some additional concerns related to latency and overhead. As information arriving at the RRH and/or BBU needs to be inserted in an Ethernet frame, header-related overhead is introduced per frame. To ensure that this overhead is small, and does not waste the potential bandwidth gains from baseband functional splitting, it would thus be desirable to fill an Ethernet payload, before sending a frame. However, waiting to fill a payload, introduces additional latency, possibly using up some of the stringent 4ms latency budget, explained earlier. Hence, it is important to consider the impact of packetization on the FH bandwidth and latency performance, *in conjunction with possible functional splits between RRH and BBUs*, in order to understand the feasibility and potential gains of different approaches. Summarizing, the main contributions of this paper along this direction are the following:

- 1) We first study the impact of different functional splits on *peak rates*, in order to better understand the raw (minimum) performance gains achievable by simply keeping some baseband functionality at the RRH;
- 2) We then study the joint impact of different packetization techniques and flexible centralization on latency and bandwidth, considering both peak rates as well as realistic rate statistics;
- 3) Finally, we use our results to identify desirable joint packetization-split options, and provide some initial insights on the potential gains achievable in terms of RRH traffic multiplexing;

To our best knowledge, this is the first paper to jointly study packetization and functional splits on the FH network.

The rest of this paper is organized as follows. In Sec. II we introduce the problem setup, discussing the FH network architecture and possible functional splits considered. Sec. III presents an initial peak rate analysis over the constrained FH link, in order to derive some initial results that will be necessary to further study the packetization impact on FH capacity. Sec. IV focuses on the packetization process itself. Sec. V considers the user impact on multiplexing, and Sec. VI then provides simulation results to validate the joint impact of packetization and baseband processing splits. Finally, Sec. VII discusses some related works and Sec. VIII concludes the paper.

II. PROBLEM SETUP

A. C-RAN Topology

When the C-RAN concept first appeared, a single direct FH link was assumed to connect an RRH to the BBU cloud. However, due to concerns related to scalability, cost, and multiplexing, it is expected that the FH will evolve towards more complex, shared topologies, similar to the backhaul network [8]. In this paper, we focus our discussion around a very simple topology, which is however characteristic of

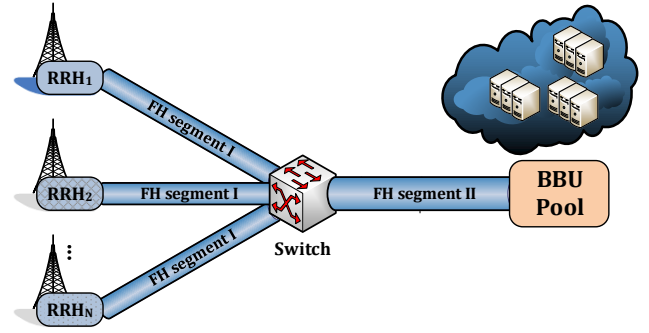


Fig. 2: Considered C-RAN topology

this envisioned evolution, presented in Fig. 2. Without loss of generality, we assume a capacity of 4 Gbps for FH segment I and a capacity of 20 Gbps for FH segment II, in order to be able to quantify the different tradeoffs.

B. Split over uplink functions

One main goal of the paper is to understand how splitting baseband functionality between the RRH and BBU will affect the experienced FH rate and latency. Fig. 3 presents the baseband processing chain, and five possible different functional splits. Without loss of generality, we focus on the uplink.

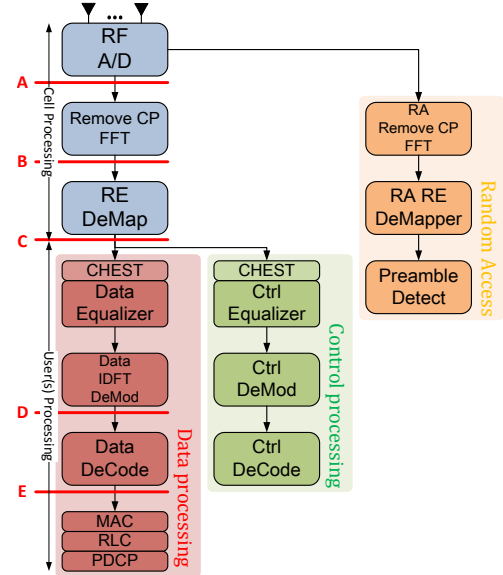


Fig. 3: Possible splits on LTE uplink function

- **Split A:** The RRH includes only time-domain RF and A/D while the BBU includes all other functions. This is the standard split considered in C-RAN.
- **Split B:** The RRH further removes the cyclic prefix (CP) in the time domain. It then applies DFT, transforming samples from the time to the frequency domain, and finally removes guard band sub-carriers.
- **Split C:** Includes also the resource element demapper in the RRH, which categorizes used resources based on pre-allocated uplink information of each served UE. In this split, essentially per-cell processing takes place inside the RRH, and per-UE processing in the BBU.

- **Split D:** The RRHs now also do some per-user processing. Channel estimation based on control and data reference signal (RS) symbols is performed and the estimation result is then applied for equalization. Afterwards, it applies IDFT and demodulation outputs the log-likelihood ratio (LLR) of each bit to BBU.
- **Split E:** The RRHs further perform bit-rate processing, including de-scrambling, de-rate matching, channel decoding and CRC check. The BBU will receive whole transport block in bits for higher layer processing.

We believe that these five splits represent the most reasonable functional split points, and that going beyond split E is not meaningful, as very little benefits from centralization would remain [3]. In the next section, we perform an initial study assuming that each cell operates at its *peak rate*, in order to acquire an initial understanding on the potential FH rate reduction by different splits, and the number of RRHs that could be supported by the Ethernet switch and aggregating FH segment 2 (without assuming any *statistical* multiplexing). Subsequently, we will move beyond peak rates, and more carefully considered the interplay between packetization and inter-cell rate variability.

III. PEAK RATE ANALYSIS

In this section, we will first perform a peak rate analysis, assuming the maximum uplink/downlink data rate that can be supported by a BS, which will provide us with an initial understanding of the impact of different splits on FH rates and its relation to some key parameters such as, number of carriers, MIMO layers, modulation order, and transport block size. We do not yet consider the packetization impact on FH latency and rates. Specifically, we focus on three LTE/LTE-A cell configurations as shown in TABLE I. For simplicity, we do not consider the number of sectors in a cell, which could be introduced as a constant factor in the calculations.

TABLE I: Configuration parameters per scenario

Scenario	1	2	3
Bandwidth	20 MHz		
Oversampling Ratio	1		
Rx Antennas	4		
Cyclic prefix length	Normal		
MIMO	4 Layer		
PUCCH RB	4		
SRS BW Config	7		
SRS SF Config	9		
Control Overhead	4.3%		
RA Config	0		
RA Overhead	0.3%		
Modulation	64 QAM	16 QAM	QPSK
TBS index	26	16	9
Time sample bitwidth	16		
Frequency sample bitwidth	16		
LLR bitwidth	8		

The resulting FH data rates for each of these scenarios and respective splits can be found in TABLE II. These results are computed by considering the data channel, control channel and random access channel data traffic, according to the specifications [9], [10].

One can observe that the data rate when moving from split A to B is almost halved, due to guard band and CP removal.

Continuing, split C offers little additional gains, compared to B. This is reasonable, as the rate for split C depends on the utilized resource block ratio, and since we assumed peak rates, almost all resources are used. For split D, the data rate varies, based on the applied modulation order. When the MCS order is large, Split D in fact has a negative impact, due to more bits being required to represent a sample, in 64QAM case. Finally, for split E, the data rate is determined by the sum of all transport block sizes, and it exhibits more than 90% reduction in the FH data rate, compared with split A. Nevertheless, as explained earlier, this comes at the cost of requiring all L1 processing to be performed at the RRHs.

Based on the required FH data rate, we also derive an initial figure on the number of supported RRHs under the FH capacity constraint of Sec. II-A. These results are shown in TABLE II. Given that they are based on peak rates, they offer a rather conservative estimate of potential multiplexing gains. Nevertheless, in the case of split E, the potential gains are already clearly visible.

TABLE II: Required FH data rate and maximum RRH number

Scenario	1	2	3
Split A	3.93 Gbps 5 RRHs		
Split B	2.15 Gbps 9 RRHs		
Split C	2.14 Gbps 9 RRHs		
Split D	2.63 Gbps 7 RRHs	1.76 Gbps 11 RRHs	878.3 Mbps 22 RRHs
Split E	300.8 Mbps 66 RRHs	123.9 Mbps 161 RRHs	63.7 Mbps 313 RRHs

In following sections, we provide the impact of packetization and multiplexing on FH link.

IV. JOINT IMPACT OF SPLIT AND PACKETIZATION

This section analyzes the impact of the packetization process on FH delay and capacity requirements for different functional splits.

A. Packetization overview

Packetization can be defined as a process of bundling data into packets according to a predefined protocol. In a typical uplink transmission, RRH needs to packetize the received samples before being transported over the FH link to BBU, whereas BBU needs to de-packetize so as to fetch all required samples before processing. Three metrics are important when evaluating the packetization process, namely:

- 1) **(De-)Packetization latency:** the extra time the RRH/BBU has to wait to get enough samples to build a packet.
- 2) **Packetization overhead:** related to the additional control information included in frame headers, in order for the transport protocol (Ethernet) to process a packet.
- 3) **(De-)packetization complexity:** refers to the additional processing load required by RRH/BBU to handle a packet.

This paper focuses on capacity-limited FH links without any constraints on BBU capabilities, and thus the only first two metrics are relevant. Ideally, the packetization method should minimize both the latency and the overhead simultaneously. However, reducing the header overhead per frame requires waiting to fill up every frame with data, which in turn introduces additional latency that might lead to violating the 4ms deadline. Hence, it is important to study the impact

of packetization on the FH bandwidth and latency performance for different functional splits between RRH and BBUs. Fig. 4 shows the trade-off between packetization latency and overhead for a given split. The FH capacity constraint and maximum allowed latency constraint will be introduced in following subsections.

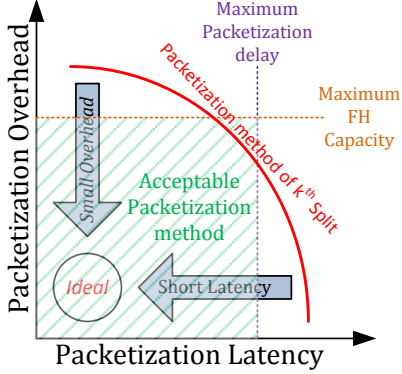


Fig. 4: Trade-off between packetization overhead and latency

B. Packetization impact on FH delay

In order to formulate the maximum allowable packetization latency, we start from the real time constraint imposed by the LTE FDD standard, i.e., the 8ms HARQ round trip time. Both the downlink HARQ ACK/NACK transmission timing should be at $(n + 4)^{th}$ subframe given that the uplink data is transmitted at n^{th} subframe [2]. We could decompose the HARQ timing constraint into the following components:

- 1) Uplink acquisition time at RRHs
- 2) Uplink processing time at RRHs and BBU
- 3) Uplink packetization latency
- 4) Uplink FH transport time
- 5) Downlink de-packetization latency
- 6) Downlink processing time at BBUs and RRH (including ACK/NACK)
- 7) Downlink FH transport time
- 8) Downlink transmission time at RRHs

Since most of the downlink data content could be prepared even before uplink reception, we consider uplink components as the bottleneck, and therefore we study the impact of packetization only on the uplink direction. Based on the results presented in [2], we assume an upper-bound of 1ms for the sum of all downlink processing, and 1ms for the uplink acquisition time. The latter is highly dependent on the design and implementation of the radio frontend and A/D converters. The maximum FH transport delay is set to $250 \mu s$ following the NGMN recommendation [11]. Moreover, the sum of BBU and RRH processing are both configuration-dependent (e.g MCS, PRB, MIMO) and platform-dependent (e.g Virtualization environment, CPU Architecture, CPU Frequency) as stated in [2]. We applied the processing time model in [2] and assume fully-parallel MIMO processing without considering an extra delay for inter-layer interference cancellation. For example, when applying the processing time model with DOCKER virtualization environment, the most time-consuming environment in [2], for 100 PRB (20MHz Bandwidth) and MCS index 27 (64QAM), the required RX processing time is $1562.3 \mu s$.

Hence, the maximum allowable packetization latency can be derived given the above-mentioned assumptions (see Fig. 4).

$$\begin{aligned}
 T_{packetization} &= (T_{HARQ} - T_{Acq} - T_{Prep}) - (T_{BBU} + T_{RRH}) - T_{trans} \\
 &= (4ms - 1ms - 1ms) - 1562.3 \mu s - 250 \mu s \\
 &= 187.7 \mu s
 \end{aligned} \tag{1}$$

C. Packetization impact on FH capacity

The available FH capacity to transport the I/Q samples depends on the packetization overhead. This overhead stems from the transport protocol, such as Ethernet headers and trailers, as well as application specific control information, such as BBU/RRH port mappings and time stamp. It is included in every packet regardless of the maximum transmission unit (MTU) and the actual payload size. Thus, the number of supported RRHs for a given FH capacity also depends on the packetization overhead, which is shown as a constraint in Fig. 4.

D. Packetization method attributes

Before specifying packetization methods of each split, we define two attributes of each packetization method:

- 1) **Packetization interval:** refers to the time at which all the received samples can be bundled to form packets. Generally, the maximum time interval is the Transmission Time Interval (TTI); however, shorter intervals specific to data and control symbols (e.g. RS symbols) are required to reduce the latency (e.g. referred as RS-aware as follows):
 - Symbol-based,
 - Slot-based, Slot-based-RS-aware,
 - Subframe-based, Subframe-based-RS-aware.
- 2) **Packetization unit:** refers to the data unit to be packed. Several units are available :
 - Sample: this is the minimum unit for packetization, which depends on the split:
 - Split A: Time-domain I/Q sample,
 - Split B & C: Frequency-domain I/Q sample,
 - Split D: LLR of each bit,
 - Split E: Bit.
 - RB: all samples of RB are packetized together,
 - UE: all samples allocated to the same UE are packetized together.

By combining the above two attributes, we could design different packetization method. Since we are focusing on the number of RRH supported over a FH link, the minimum sample-unit packetization which uses minimum overhead is considered. In next subsection, we are going to discuss the applicable packetization interval for each split.

E. Packetization method of each split

1) *Split A & B:* In splits A & B, RRH and BBU operate on raw symbols, and therefore the packetization latency is the time difference between the reception and transmission of each symbol. Thus, all the packetization intervals are applicable in this case.

2) *Split C*: In split C, RRH and BBU operate on UE-specific samples, and therefore the data rate is correlated with UE traffic allowing to exploit the multiplexing gain obtained from the MCS and MIMO layer variability. Because BBU performs the channel estimation based on data/control RS symbols before the equalization and demodulation processes, the packetization latency for split C is the time difference between the reception and transmission of the last data/control RS symbol. Similarly, all the packetization intervals are available for this split. Furthermore, in order to reduce the latency of the last control/data RS symbol, packetization of the last data/control RS symbols can be done immediately after its reception, allowing slot-based-RS-aware or subframe-based-RS-aware packetization methods.

3) *Split D*: For split D, RRH and BBU operate on UE-specific LLRs allowing packetization to be performed only on per slot or subframe basis.

4) *Split E*: For split E, both RRH and BBU operate on subframe basis. Thus the packetization can only be done on per subframe basis.

TABLE III shows all the available packetization methods for each split.

TABLE III: Packetization method of each split

	Symbol-based	Slot-based	Slot-based-RS-aware	Subframe-based	Subframe-based-RS-aware
Sample-level	Split A,B,C	Split A,B,C	Split C	Split A,B,C,D,E	Split C

V. IMPACT OF MULTIPLEXING

In Sec. III, we presented a preliminary number of supported RRHs without considering the potential multiplexing opportunities. In practice, when several RRHs experience different spectral efficiency due to individual channel qualities per UE, it is not possible for all RRHs to always have peak load. Thus, considering multi-cell multiplexing impact, more RRHs can be supported for those splits exploiting the UE-specific processing.

Several works try to model the multiplexing impact, for example, China mobile provides daily load measurement on six cells, and suggests to use three levels (e.g. idle, medium and busy) to categorize daily load [7] but without formulating specific bounds. NGMN provides two levels of load (i.e. busy and quiet hours) and forms two bounds: Lower Provisioning Bound (LPB) and Conservative Lower Bound (CLB) for back-haul provisioning for multiple eNBs [12] but not considering transient region of the two loads. In this section, we provide the average and 95th percentile data rate (i.e. data rate is below this value for 95% time) for different UE density based on the system-level simulation results shown in Fig. 5.

From Fig. 5, we observe the 95th percentile data rate belonging to different user density could be divided into three distinct regions:

- Low Density (LD) (i.e. Density $\in [1, 2]$): $R_{LD}^{95^{th}}$
- Medium Density (MD) (i.e. Density $\in (2, 10]$): $R_{MD}^{95^{th}}$
- High Density (HD) (i.e. Density $\in (10, 30]$): $R_{HD}^{95^{th}}$

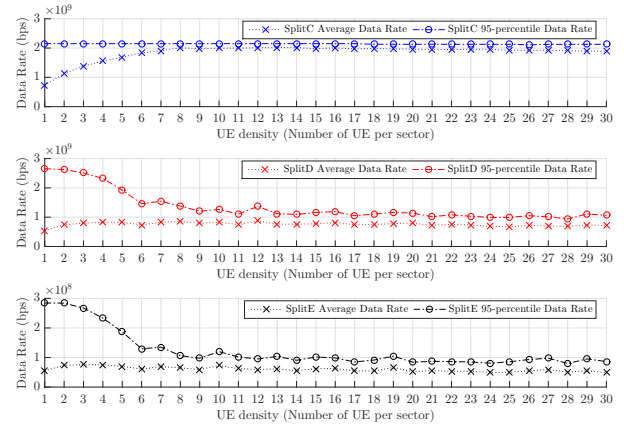


Fig. 5: Data Rate for different UE density

Then two bounds for N cells are proposed applying the three regions: Aggregated Common Bound (ACB) in (2) and Aggregated Strict Bound (ASB) in (3). Both bounds refer to the hexagonal cell geometry in which the central cell is in low density, cells of second ring are in medium density.

$$ACB = R_{LD}^{95^{th}} + 6 \cdot R_{MD}^{95^{th}} + (N - 7) \cdot R_{HD}^{95^{th}} \quad (2)$$

$$ASB = \frac{N}{19} \cdot \left(R_{LD}^{95^{th}} + 6 \cdot R_{MD}^{95^{th}} + 12 \cdot R_{HD}^{95^{th}} \right) \quad (3)$$

Fig. 6 illustrates the different bounds considering $N=19$ hexagonal cell planning with $N \cdot R_{HD}^{95^{th}} > R_{LD}^{95^{th}} > R_{HD}^{95^{th}}$. The LPB is formed only from high density cells and CLB further extends the bound to include one cell in low density but rest are in high density. As for ACB, it further considers six cells belongs to second ring are in medium density. The low density, medium density, high density cells of ASB follow distribution: $\left[\frac{1}{19}, \frac{6}{19}, \frac{12}{19} \right]$.

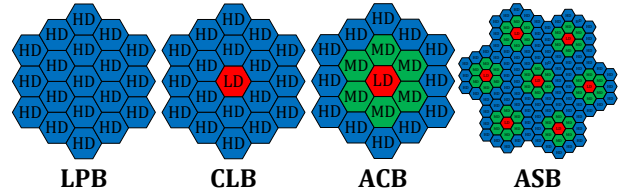


Fig. 6: Different bounds on hexagonal cell planning

VI. SIMULATION RESULTS

In this section we present the main numerical results and discuss underlying insights on the maximum number of supported RRH over a capacity-limited FH considering both packetization and multiplexing gain. Most of the simulation parameters applied to UE, RRH, and BBU are taken from 3GPP standards (TS25.942, TS36.942, TS36.814) and NGMN documents [13]. To provide a fair comparison with the peak rate analysis, the same RRH setting is used as described in Section III scenario 1, which supports 64QAM and 4-layer MIMO. UEs follow full-buffer traffic model and are moving following random walk mobility model. The detail simulation parameters are listed in TABLE IV.

Fig. 7 depicts the simulation flow in which the parameters in TABLE IV are applied in both initialization and channel

TABLE IV: Simulation parameters

Carrier frequency	2.0GHz
System Bandwidth	20MHz
Cyclic Prefix	Normal CP length
Uplink Tx/Rx Antennas	4 Tx antenna/4 Rx antenna
Uplink transmission mode	2
Inter-site distance	500 meter
Pathloss model	Urban model
Thermal noise density	-174dBm/Hz
Minimum coupling loss	70dB
Shadowing mean	0
Shadowing std	8
Inter-cell shadow correlation	0.5
Inter-sector shadow correlation	1.0
Propagation channel model	EPA
Initial UE distribution	Uniform distribution
UE speed	Randomly selected from [3, 30, 120]km/hr
UE direction	Uniform distributed in [0, 360] degree
UE antenna gain	0dB
Cell antenna pattern	3-sector cell pattern
Cell antenna gain	15dB

model step. The simulation is then executed and all processed samples by RRHs are then packetized for FH transmission.

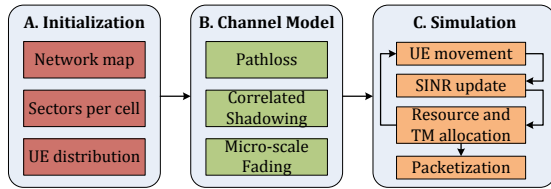


Fig. 7: Simulation flow

We use the FH capacity constraint, described in Section II-A, as well as the FH delay constraint, obtained from equation 1, to identify the applicable packetization methods. Note that the $187.7\mu s$ FH delay constraint can be interpreted as 2.63 symbol duration in case of normal CP.

A. Split A & B

Simulation results of split A and B are summarized in TABLE V for two different Ethernet packet format, standard MTU (1500 bytes) and jumbo MTU (9000 bytes). It can be seen that the data rate is constant for all the UE densities. This is because both splits operate on the cell-level samples, which has no variability. Therefore, no multiplexing gain is achievable for split A and B. We note that the 95th percentile packetization latency does not violate the FH delay constraint for different packetization methods, whereas the 95th percentile packetized data rate does violate the FH capacity constraint of 4 Gbps with the standard MTU.

TABLE V: Simulation results for split A and B

Split	Frame format	Packetize interval	Data Rate (bps)	Packetize latency (Symbol)
			95 th percentile	95 th percentile
A	Standard	Symbol	4.1506E+09	0
		Slot	4.1431E+09	1
		Subframe	4.1424E+09	1
	Jumbo	Symbol	3.9671E+09	0
		Slot	3.9671E+09	1
		Subframe	3.9665E+09	1
B	Standard	Symbol	2.2727E+09	0
		Slot	2.2665E+09	1
		Subframe	2.2658E+09	1
	Jumbo	Symbol	2.1766E+09	0
		Slot	2.1691E+09	1
		Subframe	2.1691E+09	1

Among all the packetization methods satisfying both constraints, subframe-based method achieves the minimum FH data rate for both frame formats and splits. Even though slot-based and subframe-based have comparable data rate for split B, due to integer number of generated packets in case of normal CP (i.e. seven symbols per slot), the data rate of subframe-based method is still lower when in extended CP case. TABLE VI presents the maximum number of supported RRH with packetization for split A and B.

TABLE VI: Maximum number of RRHs for split A and B

Split	Frame format	Packetize interval	Maximum supported RRHs
A	Standard	Any	Violates capacity constraint
	Jumbo	Subframe-based	5
B	Standard	Subframe-based	8
	Jumbo	Subframe-based	9

When comparing with the results in TABLE II, it can be seen that the standard frame format provides less number of RRHs. As expected, this is because the packetization overhead when using the standard frame format is larger than the one of the jumbo format. Based on this result, we only consider the jumbo format for the remaining splits. In addition, we note that packetization brings no benefits for split A and B, as there is no multiplexing gain.

B. Split C & D & E

TABLE VII presents the results for the packetization latency and the aggregated bound of 19 RRHs forming three concentrated rings (see Fig. 6). We note that the slot-based and subframe-based packetization intervals are not applicable for split C as they need more than two symbol duration (95th percentile) to just fill the packet, which often happens for low density cells due to a small number of allocated PRBs. Thus, they violate the delay constraints for data RS and both data and control RS symbols, respectively. By adding extra packetization intervals, namely slot-based-RS-aware and subframe-based-RS-aware, the packetization waiting time will be reduced for the RS symbols.

TABLE VIII presents the maximum number of supported RRHs with packetization for split C, D, and E, considering subframe-based-RS-aware as it has the minimum data rate. It can be observed that the results of the split C are comparable to that of peak rate analysis (c.f. TABLE II). This is because the 95th percentile satisfaction is very close to the full RB utilization as shown in Fig. 5. As for split D and E, results of 95th percentile satisfaction reveals a significant multiplexing gain compared to the peak rate analysis.

Based on the results, the maximum number of supported RRH can be significantly increased by exploiting the multiplexing gain. Fig. 8 compares the best packetization method per each split and the peak rate results in TABLE II.

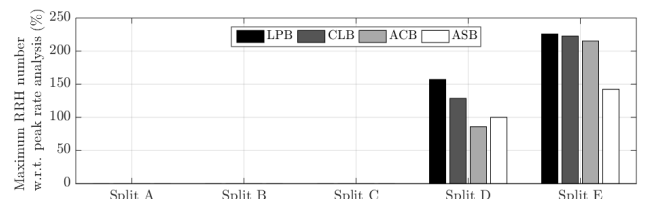


Fig. 8: Supported RRHs w.r.t peak rate analysis

TABLE VII: Simulation results for split C, D, E

Split	Packetize interval	Rate Bound for 19 RRHs (bps)		Packetize latency (Symbol)
		Bound	95 th percentile	95 th percentile
C	Symbol-based	LPB	4.1077E+10	Control RS: 0 Data RS: 0
		CLB	4.1092E+10	
		ACB	4.1180E+10	
		ASB	4.1180E+10	
	Slot-based	LPB	4.0935E+10	Control RS: 2 Data RS: 3
		CLB	4.0949E+10	
		ACB	4.1037E+10	
		ASB	4.1037E+10	
	Slot-RS-aware	LPB	4.0982E+10	Control RS: 0 Data RS: 0
		CLB	4.0997E+10	
		ACB	4.1085E+10	
		ASB	4.1085E+10	
	Subframe-based	LPB	4.0935E+10	Control RS: 7 Data RS: 8
		CLB	4.0949E+10	
		ACB	4.1037E+10	
		ASB	4.1037E+10	
Subframe-RS-aware	LPB	4.0970E+10	Control RS: 0 Data RS: 0	
	CLB	4.0985E+10		
	ACB	4.1073E+10		
	ASB	4.1073E+10		
D	Subframe-based	LPB	2.0880E+10	0
		CLB	2.2445E+10	
		ACB	2.6143E+10	
		ASB	2.6143E+10	
E	Subframe-based	LPB	1.7656E+09	0
		CLB	1.9594E+09	
		ACB	2.3684E+09	
		ASB	2.3684E+09	

TABLE VIII: Maximum number of RRHs for split C, D, E

Split	Packetize interval	Maximum supported RRHs			
		LPB	CLB	ACB	ASB
C	Subframe-based-RS-aware	9	9	9	9
D	Subframe-based	18	16	13	14
E	Subframe-based	215	213	208	160

In summary, subframe-based packetization interval proved to be the most suitable packetization method for all the splits in terms of minimum data rate. Further, RS symbol awareness can reduce the packetization latency to fulfill the delay constraint for split C. Last but not least, unlike the data channel samples, the control channel samples need to be packetize immediately for more downlink re-transmission preparation time.

VII. RELATED WORK

Recently, several standardization activities are redefining the fronthaul network towards a packet-based architecture. The goal is to design a variable rate, multipoint-to-multipoint, packet-based fronthaul interface supporting load balancing. Ref. [7] presents the Next Generation Fronthaul Interface (NGFI) and its design principles, application scenarios, and real network measurement results. IEEE 1904.3 specifies the Radio over Ethernet (RoE) encapsulations and mappings [6]. Ref. [3] and [14] describe fronthaul and backhaul requirements, transmission technologies and provide FH rate per each split under a specific configuration. Ref. [14] provides also FH data rate CDF and the number of supported BSs (or RRHs) per transmission technology with and without multiplexing gain (calculated from central limit theorem). Compared with these two works, we use the data rate as metrics to study the packetization impact, further modeled the multiplexing gain and provided numerical results under different cell loads. Ref. [10] provides a detailed peak rate analysis focusing on CPRI transmission impact, with respect to it, we highlight packetization challenge over the FH link.

Ref. [15] elaborates on per-split strategies to reduce the FH requirements while maintaining centralization advantages. Our work complement the exiting studies in that it analyzes the impact of joint packetization and split in future packet-based fronthaul network.

VIII. CONCLUSIONS AND FUTURE WORK

This work focuses on how to derive the most suitable packetization method for capacity-limited fronthaul in C-RAN architecture such that the HARQ deadlines are met. We also provide the multiplexing gain analysis for different UE densities and functional split and find the maximum number of supported RRH for the best packetization method given the split. Results reveal that there is a strong interplay between the packetization overhead and latency and that changes the fronthaul performance.

We are planning to further study the impact of queuing on the performance of fronthaul and analyze a joint packetization and split for the processing-limited BBU and RRH.

ACKNOWLEDGMENTS

Research and development leading to these results has received funding from the European Framework Program under H2020 grant agreement 671639 for the COHERENT project and from the French ANR Program under RADIS project.

REFERENCES

- [1] A. Checko *et al.*, "Cloud ran for mobile networks - a technology overview," *Communications Surveys & Tutorials, IEEE*, 2014.
- [2] N. Nikaein, "Processing radio access network functions in the Cloud: Critical issues and modeling," in *MCS 2015, 6th International Workshop on Mobile Cloud Computing and Services*, 2015.
- [3] D. Wubben *et al.*, "Benefits and impact of cloud computing on 5g signal processing: Flexible centralization through cloud-ran," *Signal Processing Magazine, IEEE*, 2014.
- [4] CPRI, "Interface specification v7.0," 2015.
- [5] OBSAI *et al.*, "Bts system reference document, v2.0," 2006.
- [6] T. Ryan, "Radio over ethernet," Feb. 8 2013, uS Patent App. 13/763,426.
- [7] Y. Zhiling *et al.*, "White paper of next generation fronthaul interface (v1.0)," China Mobile Research Institute, Alcatel-Lucent, Nokia Networks, ZTE Corporation, Broadcom Corporation, Intel China Research Center, Tech. Rep., 2015.
- [8] I. Chih-Lin *et al.*, "Recent progress on c-ran centralization and cloudification," *Access, IEEE*, vol. 2, pp. 1030–1039, 2014.
- [9] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); physical channels and modulation," 3rd Generation Partnership Project (3GPP), TR 36.211, 2015.
- [10] Small Cell Forum, "Small cell virtualization functional splits and use cases," Small Cell Forum, document 159.05.1.01, 2015.
- [11] NGMN, "Further study on critical c-ran technologies (v1.0)," The Next Generation Mobile Networks(NGMN) Alliance, Tech. Rep., 2015.
- [12] —, "Guidelines for lte backhaul traffic estimation (v0.4.2)," The Next Generation Mobile Networks(NGMN) Alliance, Tech. Rep., 2011.
- [13] —, "Next generation mobile networks radio access performance evaluation methodology," The Next Generation Mobile Networks(NGMN) Alliance, Tech. Rep., 2008.
- [14] J. Bartelt *et al.*, "Fronthaul and backhaul requirements of flexibly centralized radio access networks," *accepted for publications in Special Issue Smart Backhauling and Fronthauling for 5G Networks of the IEEE Wireless Communications Magazine*, Dec 2015.
- [15] U. Dötsch *et al.*, "Quantitative analysis of split base station processing and determination of advantageous architectures for lte," *Bell Labs Technical Journal*.